

## DOCUMENT RESUME

ED 311 878

IR 013 997

AUTHOR Stuck, M. F.  
TITLE The "How To's" of Using Word Processors and Database Managers with Qualitative Data: A Primer for Professionals.  
PUB DATE 28 Mar 89  
NOTE 43p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, March 28, 1989).  
PUB TYPE Guides - Non-Classroom Use (055) -- Reports - Descriptive (141) -- Speeches/Conference Papers (150)  
EDRS PRICE MF01/PC02 Plus Postage.  
DESCRIPTORS \*Computer Software; \*Data Analysis; \*Database Management Systems; Ethnography; Interviews; \*Microcomputers; Participant Observation; \*Qualitative Research; \*Word Processing

## ABSTRACT

This guide provides an introduction to the use of microcomputers with qualitative data. It is deliberately non-specific and rudimentary in order to be of maximum benefit to the widest possible audience of beginning microcomputer users who wish to analyze their data using software that is not specifically designed for qualitative data analysis. The tools (word processors and database managers) and the techniques for using each of these types of software are described using examples from an actual study which utilized these two types of software. The data input process is described in relation to the theoretical concerns of qualitative data analysis. Retrieval needs are also discussed. A glossary of computer-related terms and a 22-item bibliography are included. Three appendixes present coding form/design, codes and definitions, and examples of completed forms, all for PFS:FILE. (Author/GL)

\*\*\*\*\*  
\* Reproductions supplied by EDRS are the best that can be made \*  
\* from the original document. \*  
\*\*\*\*\*

U.S. DEPARTMENT OF EDUCATION  
Office of Educational Research and Improvement  
EDUCATIONAL RESOURCES INFORMATION  
CENTER (ERIC)

✓ This document has been reproduced as  
received from the person or organization  
originating it.  
□ Minor changes have been made to improve  
reproduction quality.

• Points of view or opinions stated in this docu-  
ment do not necessarily represent official  
OERI position or policy.

ED311878

THE "HOW TO'S" OF USING WORD PROCESSORS AND  
DATABASE MANAGERS WITH QUALITATIVE DATA:  
A PRIMER FOR PROFESSIONALS

Dr. M. F. Stuck  
SUNY Oswego  
Sociology/Public Justice Departments  
Mahar Hall  
Oswego, NY 13126

Presented at joint session of IACE (International  
Association for Computing in Education) and AERA (American  
Educational Research Association), March 28, 1989, San  
Francisco, CA

NOT FOR DUPLICATION OR OTHER REPRODUCTION, DISTRIBUTION, OR  
QUOTATION, WITHOUT PERMISSION OF THE AUTHOR, Dr. M.F. STUCK

BEST COPY AVAILABLE

PERMISSION TO REPRODUCE THIS  
MATERIAL HAS BEEN GRANTED BY

M. F. Stuck

TO THE EDUCATIONAL RESOURCES  
INFORMATION CENTER (ERIC)."

TR013997

THE "HOW TO'S" OF USING WORD PROCESSORS AND  
DATABASE MANAGERS WITH QUALITATIVE DATA:  
A PRIMER FOR PROFESSIONALS

PREFACE

This paper is written in response to two different experiences: (1) various comments made by participants at an ASA sponsored workshop held at Stanford University, July, 1987, entitled "Using Computers in Qualitative Research"; (2) my experience as a teacher who teaches teachers about the uses of microcomputers in education. This paper seeks to answer comments and questions raised by the individuals which were not clarified by the various segments of the workshop and inquiries made by students in my classes (Microcomputers in Education) which are not addressed fully by the various publications which exist to date dealing with microcomputer technology and qualitative data (see bibliography).

The information contained herein is deliberately non-specific with regard to machine and software and deliberately rudimentary in order to be of maximum benefit to the widest possible audience of BEGINNING users of microcomputers who seek to use software NOT specifically designed for handling of qualitative data to manipulate qualitative data.

I am aware of such programs as QUALPRO, THE ETHNOGRAPH, TAP, ASK SAM, GOPHER, ZYINDEX, and others which are, or claim to be, designed to handle qualitative data. I do not dispute that some of these might be easier methods by which to handle qualitative data, but based on the experience of the workshop and my teaching, I still feel there is a need to present some elementary information about readily available software (i.e., word processing programs, integrated packages), probably already in use, and how that software can be used to manipulate qualitative data.

Finally, this article is designed to be a "how to" article more than an article with a specific analytical or theoretical framework, of which there are plenty, although both issues are mentioned within the article, as must they be given the merging of theory, analysis, and technique necessary when using technology with qualitative data. This article is designed to be a rudimentary "nuts and bolts" article, and while to users of specially designed software packages for qualitative data or for more sophisticated computer users, this information may seem both cumbersome and terribly non-sophisticated, I still feel there is a need for this type of information. Afterall, there will always be "beginners" both on the computer and in the use of qualitative data, and there will probably always be those who do not choose, for whatever reason, or cannot afford, the software claiming to be specifically designed for use with qualitative data.

THE "HOW TO'S" OF USING WORD PROCESSORS AND  
DATABASE MANAGERS WITH QUALITATIVE DATA:  
A PRIMER FOR PROFESSIONALS

INTRODUCTION

One of the oldest methodologies in the history of American sociology is qualitative methodology, especially participant observation and in-depth interviewing, both of which date from the early Chicago School era. This methodology has been described and explained theoretically by such eminent scholars as Glaser (1965), Glaser and Strauss (1966, 1967), McCall and Simmons (1969), Spradley (1979), Hammersley (1981) and Lofland and Lofland (1984). Some of the "how to" of qualitative methodology (as well as theory) has been explained, to a degree, by scholars like Bogdan and Biklin (1982), Bogdan and Taylor (1975), Schwartz and Jacobs (1979) and Lofland and Lofland (1984). These authors have devoted a chapter or chapters to telling readers how to "sort" data through the use of index cards and photo copies of important or relevant parts of the data, and how to code data through techniques like the use of colors and shapes in the margins next to important or relevant sections of the data. These authors also explain the use of abbreviations (e.g., TN-theoretical note), ON-observational note), MN-methodological note). Most probably, all of the foregoing techniques have been used since the early days of qualitative methodology (Bogdan and Biklin, 1982).

The intent of this paper is to offer another set of "how to" techniques for dealing with qualitative data --

specifically procedures using word processors and database managers on the microcomputer for handling qualitative data. First, I present some general background information, then I will discuss "how to" techniques for word processors, then for database/file management systems: I also include a glossary of terms, and appendices with examples from a database/file management system.

### BACKGROUND

For years, the computer has been used to manipulate (i.e., "crunch") quantitative data, from relatively small amounts of data to massive data sets like NORC surveys, and to accomplish this manipulation rapidly. This linking of the computer with numeric data has perhaps led many people, including social scientists, to buy into the myth that computers are best able to deal with numbers, a belief which seems to have left qualitative social scientists in this high tech age in the "dark ages" using photocopies, colored pencils, index cards and different shaped stickers as their tools. In fact, as this paper will illustrate, the computer, or even more accurately, the microcomputer and various types of software provide excellent tools with which to handle qualitative data. However, before delving into the more specific issue of the "how to" of working with qualitative data and the computer and software of the late 1980's, several critical issues must be addressed.

The most important issue is to understand that the computer is a substitute for neither a thorough theoretical understanding of qualitative methodology nor for sound analytical processes based upon solid sociological theory and principles. The microcomputer and its software are simply tools with the capability of making research and analysis easier; they do not do the analysis -- the researcher does; and these tools must never be allowed to dictate theoretical issues or analytical categories. The researcher must keep these points in mind at all times.

The second critical factor to be kept in mind is that the microcomputer and software will not necessarily reduce the work load, and in fact, may increase it, especially in the early stages. First, there is "up front" time needed to learn what is available in terms of both hardware and software, and how to use various microcomputers and/or software packages. Second, for example, until such time as the technology is readily available (and more importantly, till sociologists can afford it), such devices as optical scanners (to transfer hand-written or typed field notes or interviews to machine readable form) will be dreams rather than reality. Thus, the researcher still must input her/his data to the microcomputer by keyboarding (typing) it in -- a time consuming, tedious process. Third, the researcher still must read the data, and must become familiar with them. These three tasks all involve time; the second and

third are not new, the first is. However, once these are accomplished, if the researcher is a competent sociologist or social scientist with a good grasp of the relevant theoretical concerns and analytical schemes then the benefits, including analytical speed, of the microcomputer may be realized.

The major focus of this paper is upon the "how to" techniques of using the high tech tools, microcomputer and software, for purposes of analysis of qualitative data. A number of authors (Agar, 1983; Becker, 1985, 1986; Becker et al., 1984; Besser et al., 1987; Brent, 1984; Friedheim, 1984; Gassaway et al., 1984; Gerson, 1984; Seidel and Clark, 1984) have addressed different aspects -- e.g., advantages and disadvantages; component parts and what they might be able to do -- of various software types and/or machine specific packages. Few, however, explain exactly "how to" do the operations and those that do are usually software- and/or machine-specific. This paper is an attempt to be quite explicit about how to engage in a number of operations via generalized kinds of software available on microcomputers.

#### TOOLS AND TECHNIQUES

One of the benefits of using microcomputer software (albeit some types more than others) is that they foster and facilitate the constant comparison (Glaser, 1965) so critical to qualitative research. The researcher is able to

look and relook at the data from various vantage points without having, for example, to shuffle and reshuffle index cards or multiple xerox copies. This ability to constantly compare electronically rather than manually should be conducive to better, more thorough analysis because it will be less cumbersome than traditional means.

Several different types of software will be described, along with an explanation and example of how to use each with qualitative data. The use of any software is predicated upon at least minimal familiarity with the data (as mentioned earlier) and some idea of what one wants to do with it. Further, for maximum utilization of the tool and for most scholarly and thorough research, all assume sound understanding of the research question(s) and issues of the methodological techniques and theoretical assumptions guiding the study at its outset.<sup>1</sup>

#### Word Processors<sup>2,3</sup>

In one sense, one of the simplest computer tools, i.e., software packages, to utilize is a WORD PROCESSOR with a SEARCH AND FIND capability. Operating on the assumption that one's data are already in a word processor FILE (see note #2), the "how to" of using the word processor and its search and find capability can be relatively simple.<sup>4,5</sup> After reading through part (or all, if it is not too long) of a file of data, one way to use the word processor as a way to assist with analysis is to go through the file

inserting key words, (i.e., actual words from the document which seem to be important\*) at the beginning of a section (e.g., paragraph[s]) where these words occur. In order to facilitate (i.e., speed up) later searching, these key words should be enclosed within DELIMITERS (e.g., less than < and greater than > signs). There can be as few or as many key words as the researcher deems important for each file (or section of the file).

Once an entire file (or set of files or entire data set)<sup>7</sup> has had key words inserted, the Search and Find operation<sup>8</sup> can be utilized to find all sections of a given document containing an EMBEDDED key word. For example,<sup>9</sup> in a data set comprised of interviews about adolescent drug use, some respondents talked a lot about "pot". For every passage/paragraph where the word "pot" was used, the key word "pot" would have been put within delimiters (e.g., less than < and greater than > signs, <pot>). Using the Search and Find function, the researcher would indicate that what she/he wanted searched for was <pot>. Identified on the screen (the exact mechanisms vary by word processor) would be each instance of where that key word occurred (followed by the actual text of the file to which the key word refers). If the researcher wanted, the termination delimiter (in this example, the greater than [>] sign) could be placed after the segment of the text to which the

embedded key word referred. Such a marking of the text could speed up the following procedure.

Using another function of a word processor, COPY, these passages could be copied via a CLIPBOARD to another newly created file called, for example "Pot", in which the researcher could place all passages identified by the key word <pot>.<sup>10, 11</sup> This new file (Pot) could then be read on screen (or printed if one finds HARD COPY easier to deal with), with an analytical eye towards comparison of the passages for similarities/dissimilarities (i.e., themes) along theoretical dimensions or issues of concern to the study. Two tasks have been accomplished by the word processor: First, it has facilitated data reduction: this new file "Pot" contains only segments of the data dealing with "pot", but has kept those segments in the context where they were found. Second, based upon comparison among the passages, a higher level of theoretical development/refinement has been achieved.

A second manner in which the concept of key words can be used is a bit more general. Rather than entering key words which correspond to actual words used in the original data, a more general or more abstract key word could be used, again, based upon the actual content but not the specific word.<sup>12</sup> Thus, for example, in the same data set of interviews dealing with adolescents and drug use, rather than the key word "pot" as in the earlier example, the key

word here is a more general one, "marijuana". Thus <marijuana> would be inserted as the key word prior to every passage where terms like "marijuana", "pot", "reefer", "weed", or other such names for marijuana appear. Again using Search and Find and Copy procedures a new file could be produced containing all the passages identified by the key word <marijuana>. This file potentially will be much larger, given the number of other words for pot, than the file in the earlier example comprised of only those passages identified by the key word <pot>. As in the earlier example, this file could then be read (on screen or in hard copy) for analytic purposes -- themes, similarities/dissimilarities among passages, theoretical development and/or refinement. Again, data reduction, constant comparison, and data analysis have been facilitated by the use of a word processor as a tool.

The foregoing has explained some ways to use a microcomputer word processing package to facilitate analysis of qualitative data. Again, it must be emphasized that the microcomputer and its software do not and cannot do the analysis for the researcher. They can aid in certain techniques of qualitative methodology: Data reduction, constant comparative analysis, theory development and refinement.

File Managers/DataBase Management Systems

A second type of microcomputer tool which can be utilized for analysis of qualitative data is a traditional file manager or database management system.<sup>14</sup> As is the case with all software, depending on the sophistication of the product it will be more or less user friendly, more or less easy to learn and use. As opposed to using the word processor and key words to handle qualitative data, more preliminary work is needed for the most efficient and effective use of a file manager or database management system. In order to minimize early frustration and maximize the potential of this tool, the researcher needs a good grasp of the research questions and the theoretical issues of the study, at least a moderate familiarity with the data, and some general idea of the kinds of analytic categories into which the data might fall. In other words, to maximally use a file manager or database management system, some preliminary type of analyses may need to be done (if only in the head of the researcher).

In a standard file manager or database management system there exist several components<sup>15</sup>: the FORM or RECORD which contains all the information for a given interview, session of observation, subject -- whatever unit you decide constitutes a kind of mini-data set; the FILE, which contains all the forms in a given study<sup>16</sup>; the FIELDS, which are the categories or types of information you want

recorded; and the VALUES, which means your data (either verbatim from notes, interviews, etc., or in the form of codes). Of most concern in terms of "how to" are the form/record and its fields.

Before addressing "how to" processes, several other issues need to be noted. It is here that the up front time to learn about and learn how to use software (mentioned in the beginning of the paper) is important, and where the researcher must have some preliminary (at least) sense of direction. The researcher must be aware of whether or not the software selected, for example, allows the creation of task specific format (i.e., determining the number of fields, where they will go on the form, etc.) versus whether the form design is for all practical purposes "LIMITED". (A "limited" form, if the researcher is not careful, might dictate the type of analysis one ends up engaging in). One must also know whether or not the software allows variable length fields and how long is the maximum field length. If, for example, one thing the researcher wants to do is enter complete quotes or descriptions and field length is limited to, say, 26 characters, the entering of verbatim quotes and descriptions would not be possible.

Once the foregoing issues have been dealt with, the researcher is ready to deal with the "how to" of using the file manager or database manager. It is usual that the first step is to decide what part(s) of your data will

correspond to each form. For example, on one project, the author's data set consisted of 100 transcribed verbatim interviews, and the decision was made that each form would correspond to a different transcript. Thus, in the completed file there were 100 different forms/records.

The next step is more difficult, and requires in depth familiarity with the data: one must now define fields.<sup>14</sup> Some initial fields will seem obvious, one such is transcript number and demographic information. For example, for demographic information, each item would constitute a separate field: Age, sex, race, ethnicity, religion, grade in school. Further field development is dependent upon the foci and theoretical concerns of the study. For example, the author's focus was with drug use/non-use and sport involvement, and this required development of such fields as: Athlete (where VALUES would be entered to indicate whether or not the respondent was an athlete); Sport (where values would be entered to note specifically which sport(s) the individual engaged in); Pot (where values would be entered to indicate type of use, if any); Hard drugs (where values would be entered to indicate type of use, if any); Alcohol (where values would be entered to indicate type of use, if any); Reason (where values would be entered to correspond to a variety of types of reasons); QuoteS, QuoteH and QuoteA (where, respectively, complete quotations from the transcript could be entered which corresponded to

reasons for sport involvement, hard drug use or non-use, alcohol use or non-use). (See Appendix A for a complete Form with all fields.)

Once the researcher has developed the relevant fields a decision must be made as to where on the form these fields will be located. This is purely up to the researcher.

---

There is no hard and fast rule or recommendation about where on the form the fields should be located, except the common sense 'rule' that fields for which the researcher will search often are best located at the top of the form since this will speed up later searches. However, in attempting to locate fields on the form another question must be asked: How long a field will be needed for entry of each value? If the maximum number of characters that will be entered for a field length is two, then there is no good reason to leave space for 26 characters (some programs take care of this automatically). By knowing the maximum length of values to be entered, the researcher can maximize the number of fields that can be utilized (assuming that the software package has no preset limits). Since fields roughly correspond to large theoretical or substantive areas of interest, the more fields one can potentially have, the easier it is to develop a number of substantive areas of interest or generate a number of different theoretical possibilities, thereby making maximum use of the features of the microcomputer and its software.

Decisions about field length are related to the values which correspond to the field names. Here another set of decisions must be made. If the researcher wishes to enter whole words as the values for fields then, obviously, the field length must be large. The entering of whole words as values has both advantages and disadvantages. The major advantage is that the value is readily apparent; for example, if the field is POT and the value is "regular" (i.e., regular user) then with little explanation the researcher or anyone else could use the database. However, there are two disadvantages of using whole words as values: (1) time of data entry is slowed; (2) a lot of space is utilized in this fashion, both on the form and, perhaps more importantly, in terms of disk storage space; (3) retrieval time will be slowed because of the length of STRINGS to be searched when looking for a match.

Perhaps a more beneficial system of entering values which are not verbatim quotes is to develop a coding schema, in which a single (usually) letter or number corresponds to a descriptive value. In terms of the development of a coding schema, it is once again imperative that the researcher have a working knowledge of the data. Without such knowledge, the development of such a coding schema and, thus, the entry of values into field areas, would be impossible. For example, if one's data allowed one to discern that the relationship between adolescents and pot

could be described as "regular user", "occasional user", or "non-user", these would be the potential values for the field. Thus, for the field, POT, if the value was "regular user" the code "R" could be used in place of, but corresponding to, the words "regular user" (one character versus 12 characters); "O", for "occasional user" (one versus 15 characters); "N" for "non-user" (one versus 8 characters). A single letter coding schema such as the above would have certain advantages over whole words: (1) more values could be entered in each field in less space; (2) less space is utilized on the form and in terms of storage space on the disk; (3) retrieval time will be speeded up because the length of strings to be searched will be less than if whole words had been used. (See Appendix B for an example of a coding schema; see Appendix C for examples of completed forms, i.e., forms with values entered for each field.)

Once decisions have been made about fields and values, the researcher should enter a few sample forms into the system and then attempt to RETRIEVE various types of data. For this trial system test, types of data should be selected for which the researcher knows the retrieval results -- fields, values, and retrieval combinations. Such a test will often point out overlooked items, such as the necessity for an additional field (as when this researcher did a test of initial data entry and discovered that a field was needed

for indicating the respondent's grade in school); or additional values (e.g., this researcher also discovered the need for codes for "former" to place in the field "Athlete", rather than simply "organized", "recreational", or "none" as was originally thought). Such a test may also reveal that during data entry, consistent formats had not been utilized<sup>17</sup>, and, thus, retrieval could be incomplete and/or inaccurate. If, for example, one has entered birthdates and utilized the format month-day-year for some values and day-month-year for other values in the same field, the retrieval will omit all those values whose format does not correspond to the format specified in the retrieval instructions. Once such a test is complete, and if it shows that the retrieval does what the researcher wants it to do, then the researcher can enter the rest of the data for the study.

If the researcher has a working knowledge of the targeted data, the actual process of entering data accomplishes, for the researcher, the tasks of constant comparison and data reduction, unless a data entry person is hired to do this. Once the data have been reduced to codes and entered into a database manager or file manager, it is possible to put into storage the original documents whence the data were derived. If coding and taking exemplary quotations has been well-done, there should be little or no need to handle the original documents again during the

course of the study (assuming, of course, that the researcher does not change the focus of the study; in that case, the process would probably have to be engaged in all over again).

Once data have been entered, then the full power of the database or file manager comes into play. The advantages of the microcomputer over, say, index cards, photocopies become obvious. Depending on the features of the database or file manager, the researcher will be able to retrieve on any number of combinations of fields.<sup>18</sup> For example, the software<sup>19</sup> this researcher utilized allowed retrieval on as many or as few fields as wanted; thus, retrieval was done on combinations like: Age: 17, Sex: M, Race: B, Athlete: O, Sport: F, Pot: R, Reasons: , QuoteP. Such a combination would not only print to screen (or hard copy if desired) all the information for all 17 year old Black males who were participants in organized football and regular pot users but would also print out the codes for the reasons for pot use as well as any quotations pertaining to pot use. This particular software also counted the number of forms which corresponded to this retrieval specification so that the researcher not only had the printout from which similarities/dissimilarities in themes could be derived but also a count of the number of respondents who fit the particular retrieval specifications. The original form is left intact on disk, ready for subsequent retrievals --

quickly, without destruction of the original, without the need for multiple photocopies or index card.

#### CONCLUSION

The foregoing has been a deliberately simplified description of a number of "how to's" for using the power of the microcomputer with qualitative data. For the audience for whom this paper was intended, this work is meant to be but a beginning, a way to get started. Without this knowledge, the "how to" of simple usage, and the encouragement to simply get started, all the sophisticated, complicated software in the world will be of no use. To qualitative researchers using the microcomputer for the first time: good luck!

Glossary

CLIPBOARD: a temporary storage space to which segments of text can be cut and pasted for subsequent insertion into a new document; feature of word processors (and some other programs as well).

COPY: a function of a word processor, data base or file manager in which a segment of the text can be reproduced either within the document/file on which one is currently working or reproduced to a clipboard for insertion into another document/file;

in a data base manager also, a function which allows the reproduction of a blank form.

CUT AND PASTE: a function of a word processor by which a segment of the text can be removed from one position in the original document, then inserted into another position in the original document/file;

or a segment of text can be removed from the original document, placed on a clipboard, then inserted into another document/file. (May also be called "Move".)

DATA BASE MANAGEMENT SYSTEM: software whose purpose is to organize information for storage and retrieval purposes;

the database stores information in files; each file consists of forms or records; each form/record contains fields, into which are placed values.

DELIMITER: character(s) not used in the main text of the document whose function is to enclose a string (and thus become part of the string) so that it is a unique string. For example, even though your text may contain multiple occurrences of the word "pot", you may have embedded the key word "<POT>" to indicate every place where not only the word "pot" appeared, but also the words "reefer", "smoke", "dope", etc. So, if the researcher wanted to locate not every occurrence of "pot", but the more specific occurrences as noted above, the researcher could use delimiters.

Delimiters are used to help speed up searches and retrievals. E.g., in "<POT>", the less than sign (<) and the greater than sign (>) are delimiters which make the string "<POT>" unique and different from simply "POT" or "pot".

EMBEDDED: in a word processed document, strings which are not part of the original text, but which are inserted by the researcher as a way to mark, code, identify certain segments of the text for specific purposes.

**FIELDS:** in a data base or file manager, the categories or types of information you want recorded. E.g., age, sex, transcript number, quotations about pot smoking, etc.

**FILE:** in a word processor refers to an individual piece of work/document (e.g., memo, letter, book chapter, etc.) which the user has created and saved;

in a data base or file management program refers to the collection of all the forms in a given study which are structured the same way.

**FORM/RECORD:** in a data base, contains all the information for a given interview, session of observation, subject -- whatever unit you decide constitutes a kind of mini-data set.

**HARD COPY:** documents, files or portions of documents, files actually printed out on paper.

**LIMITED FORM:** a record/form in a file or database manager whose number of fields is predetermined by the software, or the number of characters for a given value space is predetermined by the software, or the arrangement of the fields is predetermined by the software -- all of which allow little or no flexibility to be decided by the user. (For example, the data base component of Appleworks limits both the number of fields one can have and the number of characters in the value portion of the field.)

**RETRIEVE:** a function of a data base or file management system which allows the user to designate certain fields or values within fields which the user wants searched for and displayed to screen or printed out.

**SEARCH AND FIND:** an operation within a word processor (sometimes other applications as well) which allows the designation by the user of a "target" string which the program then finds by searching through the document. Each occurrence of the target string will be displayed on the screen usually with the first letter or entire string highlighted or shown in inverse (depends on the particular program).

**STRING:** a group of alphabetic or alphanumeric characters, distinguished from other strings by spaces before and after them. Words are strings, a phone number is a string, a birth date (9-21-45) is a string.

**VALUES:** the information (the data -- either verbatim from notes, interviews, etc. or in the form of codes) which is placed in the fields in a data base or file manager.

**WORD PROCESSOR:** software which allows the electronic creation, editing, and manipulation of textual data.

Endnotes

1. While one of the most important functions of qualitative methodology is the development of grounded theory (Glaser and Strauss, 1967) and the constant definition and redefinition of theory possible because of the constant comparative method (Glaser, 1965), all studies have (or ought to have) some theoretical assumptions no matter how broad at the outset.

2. The tool (i.e., the microcomputer) and its software (here, the word processor) can be used for simply entering and storing data, without any attempt at further manipulation. The entering of data can be a relatively simple, mechanical procedure and will not be dealt with specifically.

3. In sections that follow, computer terms (capitalized) will be used. Most of these are defined in the glossary at the end of this article.

4. If one is initially keyboarding one's own data into the computer, the following techniques can be incorporated into the initial keyboarding, thus possibly saving time.

However, first entering data, then utilizing the following techniques, in a sense, forces a constant comparison of the data. The more times one reads the data, the more familiar with it one becomes and the more connections/comparisons one can theoretically make.

5. One or two descriptions of the "how to" of using this tool will be given as examples of the possible "how to's". The researcher may/will develop others more suitable to her/his own tasks.
6. "Importance" is a distinction the researcher makes based upon things like knowledge of the data and theoretical assumptions.
7. It is usually a good idea NOT to enter key words for an entire data set without testing the procedure to see if it works, that is, does what the user means for it to do. (As is the case with any and all uses of the computer, it can only do what the operator/user tells it to do -- the old adage about "garbage in, garbage out" is accurate -- and perhaps more so with qualitative than quantitative data.)
8. The mechanism of exactly how to use Search and Find will vary slightly from word processor to word processor, thus one must understand how to use one's own word processor package(s) and its functions.
9. All examples in this paper are real examples from an actual study in which the researcher engaged. While the data here were transcripts of interviews, the procedures are applicable to field notes and other textual data.

10. Copy, as opposed to CUT and FASTE, leaves intact the original document.

11. While the process of copying to a clipboard and then to a new file can be relatively time consuming, it is much less time consuming than reading, xeroxing, cutting up, filing then reading sections of a transcript or field notes, or reading, writing on index cards, sorting, filing, then reading selected passages.

12. Another more general keyword conception could be based upon theoretical areas of interest like "soft drug use", as opposed to "hard drug use" -- each of which may have different theoretical explanations and therefore each of which would be coded differently for transfer to distinct new files.

13. Depending on the size of your study, your software and hardware, you may need more than one file to hold all the forms/records of your project. However, files may be chained together for searching.

14. A data base management system and a file management system are not quite the same. The former usually is more sophisticated than the latter.

15. What is described here is a standard file manager and/or database manager. There are what some call to be more sophisticated "free form" database managers which utilize few of the components described here. The more "old-fashioned", standard database manager is being presented here because first of all, they are right now the conventional model for many users (since they are the types that exist in many integrated software packages, e.g., Appleworks, Microsoftworks), and secondly, the components of the standard data base may be an easier way to conceptualize how a database manager can work than would be demonstrated through a description of the more "sophisticated" types.

A problem, for some individuals, with using traditional database managers is that they require the conversion of original continuous text to discontinuous categories or fields, thus possibly, though not necessarily, doing away with the context from which came the data.

16. All decisions about and designs of fields and forms should be done on PAPER before attempting to utilize the microcomputer, as this will save much time and aggravation!

17. Consistency in entry of values may not be necessary with all data base managers and file managers in order for accurate retrieval to be accomplished. For example, some data base managers and file managers allow the specification

of partial strings (i.e., "...characters..", which means search for the designated characters without regard for characters which precede or follow) or "wild cards" (i.e., search for anything in a specified field) for retrieval purposes.

18. Some database managers will allow retrieval on as few as three (3) fields simultaneously, thus limiting the combinations of retrieval that can be done. However, with careful thought and/or the creation of sub-files it may be possible to achieve the equivalent of multiple field retrievals.

19. The software utilized was PFS: File, technically a file manager; available for the Apple II series and for the MS-DOS types of machines.

BIBLIOGRAPHY

- Agar, M. (1983). "Microcomputers as field tools". Computers and the Humanities 17:19-26.
- Becker, H.S. (1985). "Software for sociologists: Finding facts and mastering data". Contemporary Sociology 14:450-51.
- \_\_\_\_\_. (1986). "Analyzing text with computers". Paper prepared for an ASA Workshop, 6/14/86, and distributed at the ASA sponsored workshop Using Computers in Qualitative Research held at Stanford University, Stanford, California, July 9-11, 1987.
- Becker, H.S., Gordon, A.C., & LeBailly, R.K. (1984). "Field work with the computer: Criteria for assessing systems". Qualitative Sociology 7(1/2): 16-33.
- Besser, H., Yamashita, R. & Duster, T. "New computer technologies and social science research methods". Paper given at International Conference on Data Bases in the Humanities and Social Sciences, Montgomery, Alabama, July 11-13, 1987, and distributed at the ASA sponsored workshop Using Computers in Qualitative Research held at Stanford University, Stanford, California, July 9-11, 1987.
- Bogdan, R. & Biklin, S.K.. (1982). Qualitative Research for Education: An Introduction to Theory and Methods. Boston: Allyn and Bacon.
- Bogdan, R. & Taylor, S. (1975). Introduction to Qualitative Research Methods. New York: Wiley.
- Brent, E. (1984). "Qualitative computing: approaches and issues". Qualitative Sociology 7(1/2): 34-60.
- Friedheim, E.A. (1984). "Field research and word processor files: a technical note". Qualitative Sociology 7(1/2): 90-97.
- Gassaway, B.M., Elder, W.L., & Campbell, J. (1984). "Word processors for qualitative sociologists: a review essay". Qualitative Sociology 7(1/2): 157-168.
- Gerson, E. (1984). "Qualitative research and the computer". Qualitative Sociology 7(1/2): 61-74.

- Glaser, B.G. (1965). "The constant comparative method of qualitative analysis". Social Problems 12(4): 436-445.
- Glaser, B.G. & Strauss, A.S. (1967). The Discovery of Grounded Theory. Chicago: Aldine.
- \_\_\_\_\_. (1966). "The purpose and credibility of qualitative research". Nursing Research 15(1): 56-61.
- Hammersley, M. (1981). "Using qualitative methods". Social Science Information Studies 1(4): 209-220.
- Lofland, J., & Lofland, L.H. (1984). Analyzing Social Settings: A Guide to Qualitative Observation and Analysis, 2nd Edition. Belmont, CA: Wadsworth.
- McCall, G.J., & Simmons, J.L. (1969). Issues in Participant Observation. Reading, MA: Addison-Wesley.
- Schwartz, H. & Jacobs, J. (1979). Qualitative Sociology. New York: Free Press.
- Seidel, J.V. & Clark, J.A. . (1984). "The ethnograph: a computer program for the analysis of qualitative data". Qualitative Sociology 7(1/2): 110-125.
- Software Publishing Corporation. (1983). PFS: File.
- Spradley, J. (1979). The Ethnographic Interview. New York: Holt, Rinehart and Winston.
- Stuck, M.F. (1985). Adolescents, Drugs, and Sports Involvement: A Qualitative Study. Unpublished Dissertation, Syracuse University. Syracuse, NY.

How To's of DBSM and WPs 29

APPENDIX A: CODING FORM/DESIGN  
(PFS:FILE)

How To's of DBSM and WPs 30

PFS:FILE -- Coding Form

|          |         |      |        |          |           |
|----------|---------|------|--------|----------|-----------|
| TRANS.:  | AGE:    | SEX: | GRADE: | RACE :   | RELIGION: |
| ATHLETE: | REASON: |      |        | SOURCE:  | SPORT:    |
| BEER:    | REASON: |      |        | SOURCE:  |           |
| ALCOHOL: | REASON: |      |        | SOURCE:  |           |
| HDRUGS:  | REASON: |      |        | SOURCE:  |           |
| CIGS:    | REASON: |      |        | SOURCE:  |           |
| TYPE:    | SOURCE: |      |        | QUOTES:  |           |
| TYPE:    | SOURCE: |      |        | QUOTEP:  |           |
| TYPE:    | SOURCE: |      |        | QUOTEAE: |           |
| TYPE:    | SOURCE: |      |        | QUOTEH:  |           |
| TYPE:    | SOURCE: |      |        | QUOTEC:  |           |

ATTACHMENT:

Key: Appendix A

For each of the 100 transcripts, one of the PFS:FILE forms was utilized. Each of the fields was coded using the codes in Appendix B.

**Field Headings:**

- Trans. = Transcript Number
- Athlete = Participant or non-participant
- Sport = Specific sport(s) played
- Reason = Codes for reasons (for use, non-use of drugs; involvement, non-involvement in sport)
- Source = Codes for where/how individual got her/his reasons for the various foci of the study
  
- Type = Code for type of substance use
- QuoteS = Excerpts (quotes) about Sport involvement/non-involvement
- QuoteP = Excerpts (quotes) about marijuana (pot) use/non-use
- QuoteA = Excerpts (quotes) about alcohol (beer, wine, or liquor) use/non-use
- QuoteH = Excerpts (quotes) about hard drug use/non-use
- QuoteC = Excerpts (quotes) about cigarette use/non-use (not a focus of this study)
  
- Attachment = Page(s) for inclusion of anything which would not fit into other fields of the form

APPENDIX B: CODES AND DEFINITIONS  
(PFS:FILE)

CODING SCHEME (PFS:FILE)

The following are the codes and their definitions as used on the PFS:FILE coding form (Appendix A). These codes and their definitions were derived from the essence of the adolescents' own discussions, and were adopted as methodological conventions. For some of these codes, illustrative excerpts are given.

RACE:      W = White  
              B = Black  
              O = Other

RELIGION:    P = Protestant  
                  C = Catholic  
                  J = Jewish  
                  N = None  
                  O = Other (e.g., Jehovah's Witness)

ATHLETE:     O = Participant in Organized sport  
                  R = Participant in Recreational sport  
                  N = Non-participant in sport  
                  F = Former participant in sport

SPORT:        B = basketball                      G = gymnastics  
                  F = football                      S = soccer  
                  D = baseball/softball            TR = track  
                  T = tennis                        R = running  
                  H = handball                      SK = skating  
                  V = volleyball                    W = wrestling  
                  P = bowling                        Z = lacrosse  
                  I = golf                            SW = swimming  
                  C = bicycling                      A = skiing  
                  AA = alternate activities (i.e., non-sport)

ALL SUBSTANCES: Coded for type of user  
                      R = regular use (more than once per week, or  
                                          more than four times per month)  
                      O = occasional use (less than three times per  
                                          month)  
                      T = tried (use less than four or five times  
                                          ever)  
                      N = never, none (only sips of parents beer,  
                                          liquor)  
                      Q = quit use

REASONS FOR USE/NON-USE:

- E = enhancement -- physical (e.g., use of drugs for increased athletic abilities; "...smoking [marijuana] helps motivate or something." [#143])
- F = fun, enjoyment, pleasure. (E.g., "I want to get high because it's fun." [#003])
- H = just to "have", "do"; felt like it (e.g., drinking, smoking)
- K = to get drunk/high. (E.g., "I do [drink] to get drunk because I have more fun when I'm drunk." [#114])
- M = enhancement -- "mental" (e.g., think better, improve one's mental state)
- R = religion/religious (e.g., not drinking because of claimed religious proscription)
- AV = availability (e.g., substance was "around" so it was used or tried. "...when it's around you smoke it [pot]." [#018]).
- DL = dislike substance -- taste or effects (e.g., "Yeah, I don't like it [alcohol]. Never like how it tasted or anything, and I just didn't really want to get into it." [#006])
- FP = family practice/ritual (e.g., wine drinking with meals, at holidays)
- GF = "general" fear (i.e., subject was "afraid" but offered no further elaboration. E.g., "...And I said no ... cause didn't want none [marijuana]. I was scared." [#040])
- IL = illegal, "wrong". (E.g., "I don't think it's right [drinking] for somebody, ya know, for somebody our age..." [#0261])
- LG = life goal (e.g., to become a professional athlete)
- MF = "mental" fear (e.g., loss of control, getting in trouble generally; "...I don't think really think there's anything wrong with it [drinking], unless you can't control yourself." [019])
- NA = not available (e.g., drugs which are "scarce" like hashish)
- ND = no desire to try, to use. (E.g., "I've turned down LSD cause I don't wanna get into anything harsh." [#031]).
- NO = no opportunity to try, to use. (E.g.: Q: Every try heroin? A: No way...even if I did [have have a chance], I wouldn't do it." [#035])

- PF = physical fears, physical reasons (e.g., addiction, getting sick; "they [drugs] can be addicting real fast, and then they tell us about the withdrawal symptoms, get sick, and, I just don't want to go through that." [#013])
- PG = peer group (as users, as influence to use)
- PI = price, cost, expense
- PM = pain relief -- "mental" (e.g., mood adjustment, getting "mellow"; "Oh, they were doing a lot, they were helping me alot. I was more relaxed and easy-going." [#114])
- PR = pain relief -- physical (e.g., use of codeine to enable one to continue playing sports)
- SP = sport involvement. (E.g., "I don't do any drugs when I'm skating. That's like -- that's my serious, you know, I won, money and stuff, I like it...I don't do any kind of drugs at all when I'm skating" [#037].)

SOURCE:

- E = own experience
- O = others' experience
- C = classes, school
- M = movies (TV, theatre)
- F = family (e.g., values, beliefs, practices)
- P = personal belief

How To's of DBSM and WPs 36

APPENDIX C: EXAMPLES OF COMPLETED FORMS  
(PFS:FILE)

TRANS: 003 AGE: 14 SEX: F GRADE: 7 RACE: W RELIGION: C  
 ATHLETE: XR,Q,AAX REASON: SOURCE: SPORT: XSW,T,DX  
 POT: XRX REASON: XPG,PM,F,FPX SOURCE: XE,F,OX  
 BEER: XRX REASON: XPG,P,KX SOURCE: XE,PX  
 ALCOHOL: XIX REASON: XDL,PGX SOURCE: XF,EX  
 HDRUGS: XTX REASON: XPG,PF,DL,MF,P,D,BX SOURCE: XE,OX  
 CIGS: XRX REASON: XPGX SOURCE: XEX  
 TYPE: XAAX SOURCE: XPGX

QUOTES: AA: friends' houses, home and listen  
 to music, talk on phone, TV, get high, get drunk.

Q: Involved in any after school activities? Plays, sports or anything?

A: Nope. Q: How come? A: Cause I don't like sports. I like some of them but I don't like plays. ...[in elementary school I was] um little jock. I was in all kinds of sports. Softball, tennis, everything...I don't do anything anymore.

TYPE: XRX SOURCE: XPG,FM,FX

QUOTE: Q: Like you don't smoke when you want to be relaxed maybe or stuff like that?

A: Well yea yea. Well you want to get in a like a good mood I get high. I get in a good mood.

Q: Were you with other people that were smoking again or what?

A: I don't know they smoked all the time I didn't so that didn't have nothing to do with it. I just just I don't know, I flet like getting high again. I want to get high because it's fun.

A: All my friends just got high so I did. I never stopped since then.

TYPE: XR, TX SOURCE: XPG,P,KX

QUOTE: XE,OX

A: I drink almost every night now [beer].

Q: You drink in order to get drunk? A: That's what I drink for.

Q: Do you usually drink when you're alone or with people? A: Just people.

{tried wine-disliked taste; tried liquor-disliked taste; mother's boyfriend gets her drunk and high}

TYPE: XTX SOURCE: XPG,D,P,DX

QUOTE: HASH: No, I haven't had hash in a long time. I wish I did, but. Can't get it

around anymore, it's too expensive.

ACID: B,D,MF

SPEED: PF,DL,MF

VALIUM: P,D

LUDES: FM,D, MOTHER SUPPLIES, OX'S DOSAGE

COCAINE: P,B,D

TYPE: XRX SOURCE: XPG,PFX

QUOTE: Q: How come you started smoking more?

A: I got hooked and I smoked more. I don't know.

Q: What made you start up again?

A: Just wanted one. I started smoking again.

TRANS: 029 AGE: 15 SEX: N GRADE: RACE: W RELIGION: C  
 ATHLETE: XOX REASON: XF,FPX SOURCE: XE,F,PX SPORT: XF,D,S,HX Father was a college star

POT: XNX REASON: XPF,MF,SP,PGX SOURCE: XE,PX

BEER: XOX REASON: XB,PF,PGX SOURCE: XE,F,CX

ALCOHOL: XT-WINEX REASON: XFPX SOURCE: XE,FX

HDRUGS: XNX REASON: XPF,MF,IL,PGX SOURCE: XP,EX

CIGS: NTX REASON: XPF,SP,PGX SOURCE: XP,EX

TYPE: XOX SOURCE: XF, FPX

QUOTES: A: I love football...I do play defensive tackle for Oxford...I was on the soccer

team, the track team, um I was, working at that time, for the Empire State Games...catcher in baseball...

Q: Would you say football is your favorite sport?

A: I think handball first. That's an exotic sport. Then football and baseball.

TYPE: XPF,MF,SPX SOURCE: XE,PX

QUOTE: A: [Some players] asked me if I wanted it and I said no. I figured if I was

high I wouldn't know what I was doing and I was gonna line up against some pretty big players and I wanted to know what I was doing so if I ever got hurt or could've gotten hurt, I could've known what to do.

Q: Do you think it really affects the way the guys perform on the team?

A: Yea, it does, it shows two things. It shows how much a kid can have his life influenced by something and it shows how much other kids know what to do about it and when it gets out in the field, some kids it helps and some kids it doesn't help...It's illegal...could've messed up lives...being addicted...

TYPE: XB,PF,PGX SOURCE: XE,F,CX

QUOTE: A: Q: Do you like drinking?

A: No, not really. Ah, it's just something to do, you get bored, there's nothing to do, most of the fun in drinking beer is seeing how you can get it. ...I stayed strictly with beer because I knew it had no effect on me and I was pretty much scared to go to anything higher and probably won't go to anything higher than beer...When it came up, I did it but I don't drink as I have to, when I'm alone I don't drink anything...not alone, I think it had more to do with a peer group....I pretty much knew what it would have done to me through health. They explained some of the stuff to you in health.

TYPE: XPF,MF,ILX SOURCE: XPX

QUOTE: A: I think drugs are the worst thing. The way it messes up your body. Just you

get addicted to it, you have to have it, you use money for it, it's just wrong. That's one thing I do feel strong about, not using drugs.

Q: I could imagine being the athlete that you are.

A: I find more athletes use it than not athletes; the athletes get into it, but I couldn't just do that to me, that's torturing me.

TYPE: XPF,SP,PGX SOURCE: XP,EX

QUOTE: Q: Do you have serious reasons for not smoking?

A: Sports. When smoke gets in your lungs or something, you can't run, you're coughing all the time. You're on the field, it looks like you're gonna have a heart attack, and it doesn't look good.

Q: Do any of your friends smoke?

A: Not really, all of us are really into sports and ah, just can't smoke and do sports at the same time.

Attachment : XSP, PGX --see also p. 118-9 of transcript

42

A: [Team handball], the only times that I do play are the try-outs at the Municipal State Games, and the Municipal State Games [themselves]....I play lots of basketball Basketball's the closest you're going to get to it But um

I do run, I practice shooting with a basketball. And, lift arm weights. ...Sports, most of my time is in sports. ...I plan as a big part, to get part of a scholarship, in football.

A: The only [friend] that I did bring up with me from Elmer Under School was Eaton Astor but we, we're friends but not as close as the other ones. Eaton's the one that didn't go out for sports and that was Eaton's major goof.

Q: What kinds of things do you usually do (with your friends)?

A: We play cards...pinball and electronic games...painting the water towers.

Q: I'm surprised you didn't mention playing many sports with your friends...?

A: It is but when we left Every [school]...they stayed with soccer. I hated soccer. Um, in soccer, if someone is close to me and starts pushing on me to get the ball during the game, I hit someone just by accident and lost my cool because he kicked me--I can do that in football, use my hands and push ...